Wavelet for binocular vision modeling

Louis Gaudart and Jean Gaudart*

Aix Marseille Univ, IRD, INSERM, SESSTIM, Marseille, France

* Corresponding author.

SESSTIM UMR1252, Faculté de Médecine, 27 Bd Jean Moulin, 13385 Marseille cedex 05 France
Tel: + 33 491 791 910; Fax: + 33 491 794 013
E-mail address: jean.gaudart@univ-amu.fr

Abstract

In the present study, binocular vision properties were modeled using a single elementary wavelet. Opponent responses (ON-OFF) appeared in the first stages of the neural coding in the retina. This property was assumed to build an adequate wavelet showing a positive part (On) and a negative part (OFF). We have examined the experimental orientation and position disparity given by Bishop. We assumed that the theoretical position disparity was given by a combination of two wavelets for a given orientation disparity β. A change in β implied a change in the magnitude of one of the wavelets and consequently a change in the wavelets combination. There was a close match between the theoretical and experimental position disparity curves according to the changes in orientation disparity.

Keywords: receptive field, stereopsis, orientation disparity, position disparity, wavelet.

1. Introduction

Several previous physiological studies have shown that there are various receptive field organizations within the retina, the lateral geniculate nucleus and the cortical level [37]. The type 1 exhibits center-surround spatial and chromatic opponency. The type 2 exhibits a chromatic opponency (*e.g.* blue-yellow), without spatial opponency. The modified type 2 exhibits a center-surround spatial opponency and a chromatic opponency only in the center. The type 3 exhibits an achromatic center opposed by an achromatic surround. The double opponent cells (DOC) [1] exhibit a center-surround organization with a chromatic opponency in the center (for example R+G-) and an opposed chromatic opponency in the surround (for example R-G+).

Many attempts have been made to describe the physiological properties of the visual system with the help of various theoretical models. Rodieck [33] and then Enroth-Cugell and Robson [12] introduced a difference of two Gaussians (DOG) to model ganglion cell receptive fields. Another model was given by Marcelja [27] and Daugman [8] who introduced the Gabor function and the joint space/spatial frequency uncertainty relation. The Gabor transform provides a local description of the visual process. However, the Gabor model is inadequate in some species [35, 39, 18]. Marr [28] used the Laplacian of the Gaussian ($\nabla^2 G$) called the "Mexican hat". However, this author had outlined that this function cannot describe a type 1 chromatic receptive field [29]. Kulikowski, Marcelja, and Bishop [20] used a Gabor function to relate the receptive fields to the spatial frequency and orientation. Billock [1] reported a filtering mechanism to model the double opponent receptive field. Croner and Kaplan [6] investigated the achromatic receptive field (type 3) and contrast responses of ganglion cells using the difference of two Gaussians.

In a more general way, the previous theories were made under various hypotheses but failed to explain several visual properties simultaneously. Most problems were considered individually, in isolation from other visual process. The above theoretical studies examined directly the spatial receptive field organization. However, another possible type of model can be obtained by investigating the first stages of the visual process, from visual receptors to ganglion cells. Chromatic and achromatic signals are carried by several elementary cells as cones, horizontal, bipolar, amacrine, or ganglion cells [21]. A receptive field structure is the result of several specific connectivities of these neurons. In order to explain color vision and light adaptation, Guth [15, 16], then De Valois and De Valois [9, 10] proposed multi-stage models.

2

Using a multiresolution approach for a vision process model is attractive and remains an open question. Wavelets are mathematical functions which are localized in space and analyze the signal according to scale. A sharp or a coarse analysis is known as a multiresolution analysis. The wavelet theory can be used in two different manners. First, the wavelets can be considered as a mathematical tool to investigate a signal as an electroencephalogram one [32, 34] or an electrocardiogram one [36]. The wavelet theory is especially used to investigate digital image coding [24, 25, 17, 30]. Second, the wavelets can be considered as a mathematical tool for modeling physiological properties. Li and Atick [23] have shown the principle of efficient visual coding as implemented through multiscale representation. Bradley [3] described a model of the human visual system based on the wavelet transform.

The choice of an adequate wavelet function is always a crucial and difficult factor. To build an adequate wavelet we have to investigate the first stages of the neural coding. Numerous studies [22, 38, 7] have shown that the opponent responses (ON-OFF responses) appear in the early neural synapses. An adequate wavelet must take into account this property. Moreover, considering the binocular vision, the visual judgments of the depth produce by the disparity differences of two vertical lines play an important role [40, 26]. The studies of Bishop [2] concerning the position and orientation disparities are particularly interesting. This author showed how signals originating in each eye are combined to obtain a binocular standard receptive field. Therefore, these results have to be taken into account to build a mathematical model.

We have to address the problem from the following standpoint. Is it possible to have a single conceptual framework to provide a description of several processes?

In such a way, we built a wavelet which verified the three following conditions:

a) The theoretical curves obtained have to be similar to corresponding experimental curves ON-OFF response and Bishop data [2]).

b) The number of required theoretical coefficients have to be small and each of these have correspond to a given property of the visual system.

c) The model can be used to account other visual mechanisms.

3

2. Methods

Wavelet function

The basic characteristics of the early visual processing, inside the retina, are ON-OFF responses. There are parallel pathways through the retina, with On-channels and Off-channels [19, 40]. A receptive field is composed of an ON-region and an OFF-region. The simplest function to model an "ON-OFF" signal is the Haar function H, defined by the following relations [4]:

$$
\begin{cases}
H(x) = 0 & if \quad x < 0 \\
H(x) = 1 & if \quad 0 \leq x < 0.5 \\
H(x) = -1 & if \quad 0.5 \leq x < 1.0 \\
H(x) = 0 & if \quad x \leq 1.0
\end{cases}
\tag{1}
$$

However, the Haar function, which is a wavelet, is not an adequate model to reflect a visual processing. Indeed, this function shows angular points with a gap between the positive and the negative value, whereas the response of the visual system is continuous and regular, such as the transmission of a signal by a receptive field. A more appropriate wavelet should be determined. A continuous rising or decaying phase of a curve can be modeled by hyperbolic functions. We have defined over **R*** a function Ψ, having a positive and negative region, by:

$$
\Psi_N(x;c) = \frac{1 - \tanh^2(x)}{\tanh(c)\tanh(x) - \coth(c)\coth(x)}
\tag{2}
$$
$$
c \in R^*
$$

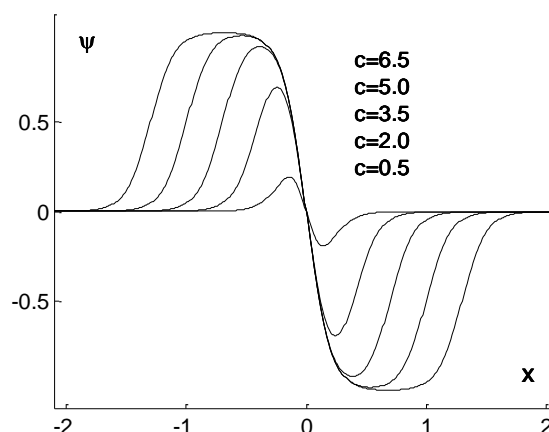**Fig. 1** Plots of the function Ψ, for five c-values. For all curves, the coefficients "a" and "z" are the same: a = 0.2; z = 0.

Such a curve can have a shape similar to a Haar curve (Fig. 1).

We have checked that the function Ψ verifies the following conditions [4].

a) *The function integral converges:* $\int_{-\infty}^{+\infty} \Psi(x)dx \;\; < \infty$

b) *The function integral is null:* $\int_{-\infty}^{+\infty} \Psi(x)dx = 0$

c) *The Fourier transform of the function is null when $\omega = 0$.*

$$\text{If } \omega = 0, \text{ then:} \quad \int_{-\infty}^{+\infty} \Psi(x)e^{-i\omega x}dx = 0$$

d) *The function has a higher number of vanishing moments* (1-order moment, 2-order moment, the even moments are null).

Consequently, the function Ψ is a wavelet, assuming a positive and a negative part. Thus, it is particularly well adapted to reflect an ON-OFF signal.

Wavelet family

The function specified by the above formula (Eq. 2) is a "mother wavelet". From the mother wavelet we have obtained a wavelet family by introducing a dilation coefficient "a" and a translation coefficient "z". Therefore, the family wavelet is given by the following formula:

$$\Psi_N(x;a,c,z) = \frac{1}{\sqrt{N}} \times \frac{1 - \tanh^2(\frac{x-z}{a})}{\tanh(c)\tanh(\frac{x-z}{a}) - \coth(c)\coth(\frac{x-z}{a})} \qquad (3)$$

$$a \in R^{+*} \quad z \in R \quad c \in R^*$$

Note that the function contains the coefficient "c" which does not form part of the classical wavelet theory. However, this coefficient plays an important role, as we will see further on. It allows taking into account some of physiological properties of the visual system, inducing the shape of the curve (Fig. 1). If "c" increases, then the curve approaches a Haar function. For a given value of "c", the theoretical coefficient "a" determines the wavelet width (Fig. 2). By varying the translation coefficient "z", one can scan the x-space (Fig. 3).
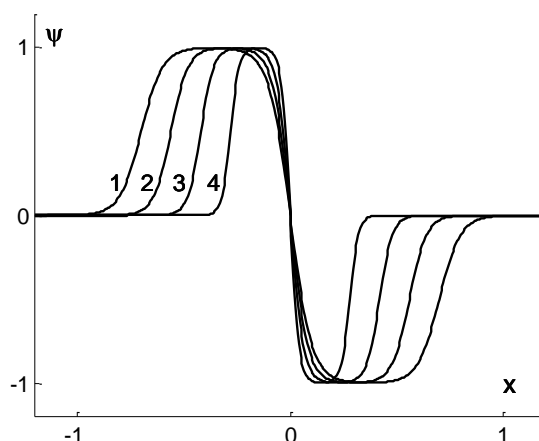

Fig. 2 Plots of the function Ψ, for four a-values. Curve 1: a = 0.10; curve 2: a = 0.08; curve 3: a = 0.06; curve 4: a = 0.04. For all curves, the coefficients c and z are the same: c = 7; z = 0.
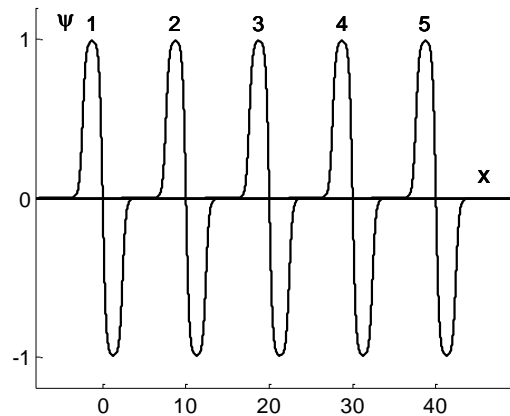
Fig. 3 Plots of the function Ψ, for five z-values. Curve 1: z = 0; curve 2: z = 10; curve 3: z = 20; curve 4: z = 30; curve 5: z = 40. For all curves, the coefficients a and c are the same: a = 0.4; c = 6.

Bishop's data

We compared the above wavelet versus the data published by Bishop [2]. The neural mechanisms have been studied by this author in the striate cortex of the cat. A stereopair of lines was viewed in a stereoscope. The two retinal images, $X_R Y_R$ (right eye) and $X_L Y_L$ (left eye) were at an angle to one another. The position disparity was deduced to the length of the line $Y_L Y_R$. The angle between the two images was the orientation disparity. The orientation and position disparity changes occurred in the right eye with the left eye image held constant. The position and orientation disparities were expressed in terms of a visual angle. The data given by Bishop [2] are represented by the points in Fig. 5a and 5b. The angle β was the orientation disparity for the right eye. The left eye image held constant at 93°. The data points were recorded in orientation steps of 3°.

7

3. Results

If we assumed that one retina response is modeled by a wavelet, then, the combination of two wavelets was able to reproduce the position and orientation disparities (two retina responses).

First, let us consider two wavelets (Fig. 4) defined by:

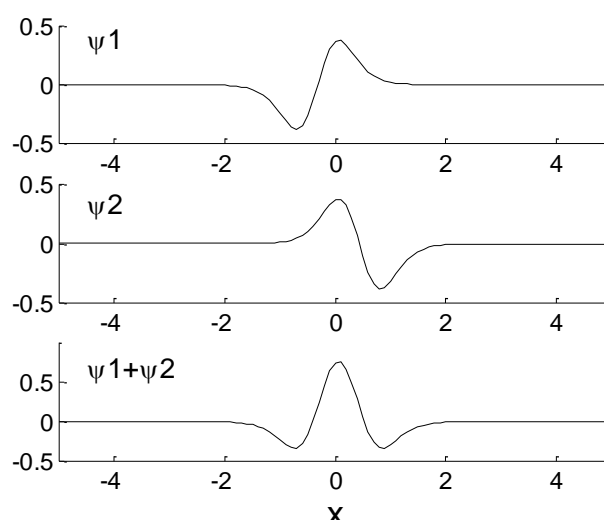$$\Psi_1(x) = -\Psi(x; a, c, z_1) \quad \text{and} \quad \Psi_2(x) = \Psi(x; a, c, z_2) \tag{4}$$

Fig. 4 The upper figure shows the wavelet $\Psi_1(x;a,c,z_1)$, the central figure shows the wavelet $\Psi_2(x;a,c,z_2)$ and the lower figure shows a binocular response. The wavelet coefficients are: $a = 0.46$; $c = 1$; $z_1 = -0.32$; $z_2 = 0.44$. The three curves share a common abscissa.

We introduced a theoretical angle α to represent the orientation disparity β and two parameters A and B. The full amplitude of the resulting theoretical curve was adjusted to the corresponding Bishop's curve by varying the parameter A and the position of the resulting theoretical curve was given by the parameter B. Thus, we obtained the following equation:

$$R(x; \alpha) = A \times (\sin(\frac{4\pi}{180}\alpha) \times \Psi_1(x) + \Psi_2(x)) + B \tag{5}$$

where x is the position disparity expressed in degree.

$R(x;\alpha)$ represents the theoretical binocular response for a given theoretical orientation disparity α.

The locations of wavelet centers are z_1 and z_2, respectively. The distance between the two centers represents the distance between the two retinas. During the experimental measurements this distance was held constant: $z_2 - z_1 =$ constant.

8

Second, the calibration of the model was achieved through a middle value of the orientation disparity: $\beta = 86°$. For this value the experimental curve was a standard receptive field. To model this curve, we have determined the values of the wavelet coefficients. The coefficient "c" was chosen in an arbitrary manner by setting: $c = 1$. The other coefficients were calculated by minimizing the normalized root mean squared deviations (NRMSD) between the data and the theoretical values. We obtained:

$a = 0.46$, $z_1 = -0.32$, $A = 360$, $B = 165$.

$$\alpha(°) = \beta - 59° \qquad (6)$$

$$z_2 = z_1 + 0.76 \qquad (7)$$

The parameter A converts an arbitrary theoretical unit into the unit of Bishop's data (a number of spikes).

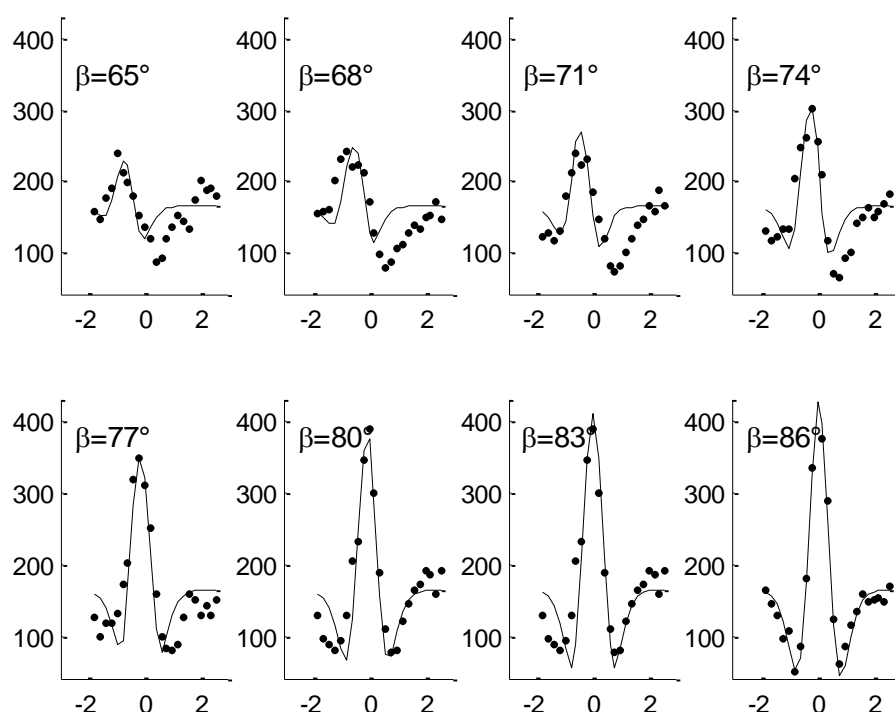By using Eq. 5 we obtained the continuous curve of the Fig. 5a for $\beta = 86°$.

Fig. 5a.
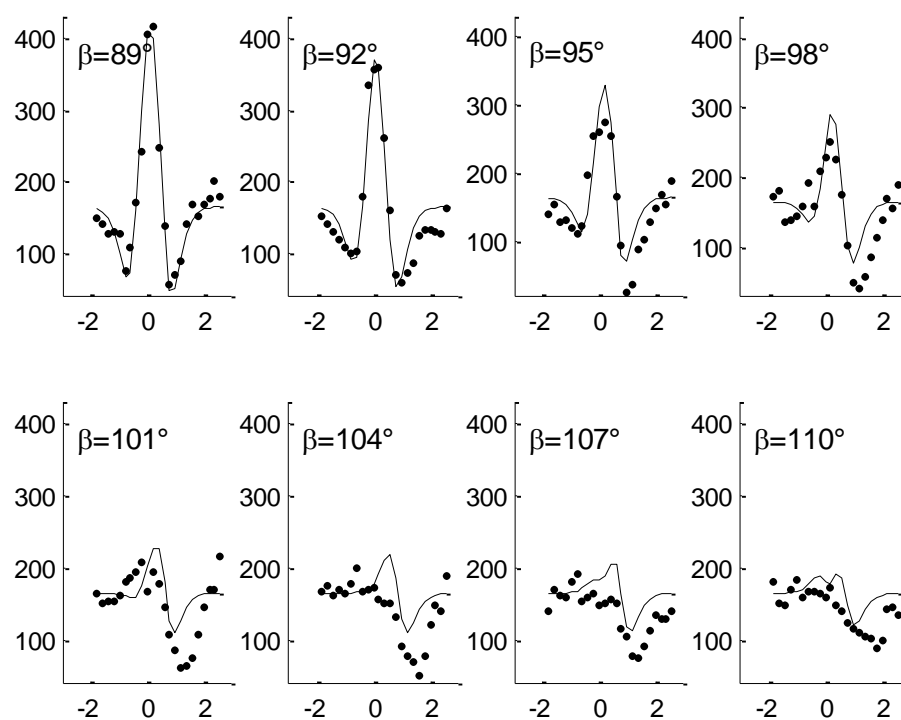
9

Fig. 5b

Fig. 5a and 5b At each orientation disparity β, a curve of position disparity is plotted (x-axis in degrees). The points represent the position disparity redrawn from Bishop [2]. The y-axis represents the total spike count. The orientation disparities β are given at increment of 3°. The continuous lines are obtained by combining two wavelets according to Eq. 5.

Third, the following results were obtained by using Eqs 5, 6 and 7, and the above values of "a", A, and B. The disparity theoretical curves for the others β-values were performed by varying α. As mentioned above, an increase in coefficient "c" produced an increase in the theoretical curves. To match the theoretical curve to the experimental points, we calculated a value of the coefficient "c" for each β-value. Following the same idea we calculated a z_1-value for each β-value.

Table 1 summarizes the performance of the coefficients calculations. The NRMSD given in this table are the minimal values which were obtained by varying "c" and "z_1".

On the basis of this analysis, the theoretical curves of the Fig. 5a and Fig. 5b were obtained. The relationship between "c" and the orientation disparity is revealed by the Fig. 6. To a first approximation, two linear relations can be regarded as satisfactory.
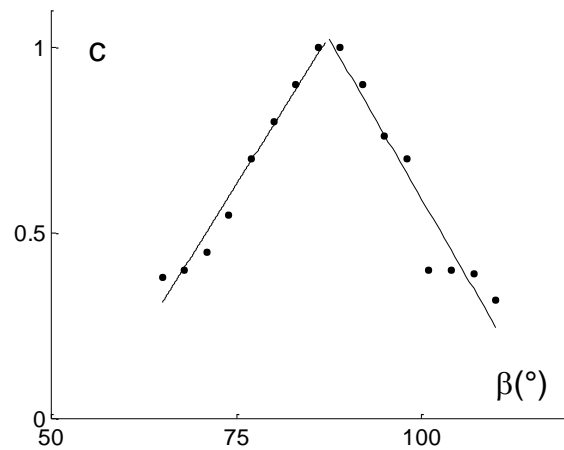
10

Fig. 6 Coefficient "c" change plotted as a function of orientation disparity β. The points were deduced from the Table 1. The straight lines were obtained by use of a linear regression.

Table 1 Wavelet coefficients "c" and "z_1" for a given orientation disparity.

β	c	z_1	NRMSD*
65°	0.35	- 1.14	0.26
68°	0.40	- 0.94	0.28
71°	0.45	- 0.84	0.23
74°	0.55	- 0.64	0.17
77°	0.70	- 0.54	0.13
80°	0.80	- 0.45	0.11
83°	0.90	- 0.38	0.06
86°	1.00	- 0.32	0.04
89°	1.00	- 0.29	0.06
92°	0.90	- 0.30	0.08
95°	0.76	- 0.21	0.13
98°	0.66	- 0.14	0.16
101°	0.40	- 0.12	0.33
104°	0.40	+ 0.11	0.36
107°	0.39	- 0.05	0.34
110°	0.32	- 0.01	0.43

* Normalized Root Mean Squared Deviation.

11

4. Discussion

We defined a wavelet function which take into account different processes of the visual information. We can obtain a plausible model by considering a certain number of visual properties.

Shape of the theoretical and experimental curves

We now have good evidence that there are ON-OFF responses in the earlier levels of the retina [31]. Consequently, we defined a wavelet which presents a positive region (ON), and a negative region (OFF). Then, we could expect that such wavelet was able to model binocular vision properties. Indeed, a good agreement was shown with the data found by Bishop [2], in particular with the position and orientation disparities. It is particularly interesting to compare the theoretical curves and the experimental points of the Fig. 5a and Fig. 5b. The theoretical changes of the curves were obtained by varying the theoretical orientation disparity α (Eq. 5). Likewise, the α-change was deduced from the β-change using Eq. 6. On the other hand, it is interesting to note that both theoretical and experimental curve changes are the same. We can see figure 5a, for $\beta = 65°$, is a unique right-hand minimum for a positive position disparity. As β (or α) increases, a second left-hand minimum appears, located in a negative position disparity. This minimum increases until the value of β reaches $86°$. As β still increases, the second minimum progressively decreases and disappears about $\beta = 104°$.

Theoretical coefficients

The required coefficients were in few in number: "c", "z", and "a". These estimated coefficients have significant physiological interpretations.

Let us recall that the coefficient "c" was not deduced from the wavelet theory. Nevertheless it has a specific meaning in the case of the binocular vision. Bishop [2] noticed that any binocular response was larger than the sum of the corresponding monocular responses. This property is reflected by the coefficient "c". The visual information processing achieved through different channels. A channel feature should be characterized by a value of "c" which is similar to an amplification coefficient. The Fig. 6 shows that the coefficient "c" is not an arbitrary coefficient. Indeed, the coefficient "c" can be deduced from the β–value.

12

The coefficient "a" determines the wavelet width. The visual system adapts to the image size. It can see a large surface or a small point. This property is reflected by the dilation coefficient. The smaller the wavelet step, the better the analysis. The Bishop's data were obtained with a stimulus width held constant. Consequently, the dilation coefficient did not play any role in the present study. The representation of an image in the visual cortex involves both spatial position and spatial frequency variables [20]. In the present study the spatial variable is the position disparity x and the spatial frequency variable is the coefficient "a" as we deduced from the Eq. 3.

The coefficient "z" is the classical translation coefficient of the wavelet theory. In the present study it indicates the spatial position of a retina with regard to the incident stimulus. If an eye position "z_1" is used as a reference value, then the second eye position "z_2" may be known. Moreover, the visual system is able to scan the viewed surface. This property is reflected by the translation coefficient. The actual eye movements explore a viewed surface using numerous successive saccades [5, 11]. Each saccade is reflected by a small variation Δz of the translation coefficient.

We can conclude that the estimated coefficients used in our study are consistent with the basic properties of the visual system.

Chromatic receptive fields and contrast sensitivity function

The present wavelet model can be used to build various chromatic fields, as we have shown in a previous study [14]. We have defined chromatic wavelets in the following way. For example, the positive segment of a chromatic wavelet represented a response of a green stimulus and the negative segment represented a response of a red stimulus. Thus, we obtained a chromatic wavelet Green-Red. Such a wavelet modeled one of the curves given by De Valois and De Valois [9, 10]. Under the same conditions we obtained a Blue-Yellow wavelet. The De Valois and De Valois [9, 10] chromatic curves were obtained by combining several chromatic wavelets. Then, the simple chromatic receptive fields (type 1 and 3, for example) were obtained by a linear combination of two chromatic wavelets. The more complex chromatic receptive fields (type 2, 2-modified, Double Opponent Cells) were obtained by combining several chromatic wavelets.

After obtaining an adequate wavelet function, it becomes possible to use all the properties of the wavelet theory. The wavelet transform is a particularly appropriate tool to model other visual system properties, such as the contrast sensitivity function, or the view of an illuminance gradient between a bright and a dark area [13]. The wavelet transform

13

provides a flexible window which automatically narrows when observing high frequency phenomena and widens when observing low frequency phenomena.

5. Conclusion

In the present approach, several properties of the visual system were modeled within a single conceptual framework. Even in view of the progress of wavelet analysis in recent years, the question is far from being settled. The place of the wavelet in the visual processing levels cannot easily be found. However, the wavelet function and the wavelet transform reflect a visual processing hierarchy. Such a representation is very economical in terms of coefficient number. We have shown that a model based on the wavelet theory enables one to find several visual properties. At the output of the retina, *i.e.* the ganglion cells, the wavelet is consistent with the measured responses. At another stage of the process, linear combinations are consistent with binocular vision properties. Finally, a wavelet transform can be situated at the cortical regions. We have explored the theory in simple visual systems with a static object, without eye movement. The model presented here is a first approach. Predictions of visual performances would require a study of various variables as time or movement. A more elaborated model should include recognition of objects at a distance.

References

1. Billock VA (1991) The relationship between simple and double opponent cells. Vision Research 31:33-42

2. Bishop PO (1978) Orientation and position disparities in stereopsis, in: Cool SJ and Smith EL (Eds) Frontiers in visual science, Springer-Verlag, New-York, USA, pp 336-350

3. Bradley AP (1999) A wavelet visible difference predictor. IEEE Transactions on image processing 8:717-730

4. Chui CK (1992) An introduction to wavelets. Chui C.K. Series, Academic Press inc., New-York, USA

5. Collins T, Doré-Mazars K (2006) Eye movement signals influence perception: evidence from the adaptation of reactive and volitional saccades. Vision Research 46:3659-3673

6. Croner LJ, Kaplan E (1995) Receptive fields of P and M ganglion cells across the primate retina. Vision Research 35:7-24

7. Dacey D, Packer OS, Diller L, Brainard D, Peterson B, Lee B (2000) Center surround receptive field structure of cone bipolar cells in primate retina. Vision Research 40:1801-1811

8. Daugman JG (1985) Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two dimensional visual cortical filters. J Opt Soc Amer 2:1160-1169

9. De Valois RL, De Valois KK (1993) A multistage color model. Vision Research 33:1053-1066

10. De Valois RL, De Valois KK (1996) On "A three-stage color model". Vision Research 36:833-836

11. Doré-Mazars K, Vergilino-Perez D, Collins T, Bohacova K, Beauvillain C (2006) The use of recurrent signals about adaptation for subsequent saccade programming depends on object structure. Brain Research 1113:153-162

12. Enroth-Cugell C, Robson JG (1966) The contrast sensitivity of retinal ganglion cells in the cat. Journal of Physiology 187:517-552

13. Gaudart L, Crébassa J, Pétrakian JP (1993) Wavelet transform in human visual channels. Applied Optics 32:4119-4127

14. Gaudart L, Crébassa J, Pétrakian JP (1996) Fonction ondelette et champs récepteurs chromatiques. ITBM 17:1-12

15. Guth SL (1991) Model for color vision and light adaptation. J Opt Soc Amer A 8:976-993

16. Guth SL (1996) Comments on « A multi-stage color model ». Vision Research 36:831-833

17. Khademi A, Krishnan S (2007) Shift-invariant discrete wavelet transform analysis for retinal image classification. Med Biol Eng Comput 45:1211-1222

18. Klein SA, Beutter B (1992) Minimizing and maximizing the joint space-spatial frequency uncertainty of Gabor-like functions: comment. J Opt Soc Amer 9:337-340

19. Kolb H (1991) Anatomical pathways for color vision in the human retina. Visual Neuroscience 7:61-74

20. Kulikowski JJ, Marcelja S, Bishop PO (1982) Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. Biol Cybern 43:187-198

21. Lee BB (1996) Minireview. Receptive field structure in the primate retina. Vision Research 36:631-644

22. Lennie P (1991) Color vision. Optics and photonics 2:10-16

23. Li Z, Atick JJ (1994) Efficient stereo coding in the multiscale representation. Network 5:157-173

24. Mallat SG (1989a) A theory for multiresolution signal decomposition: the wavelet representation. I.E.E.E. Transaction on pattern analysis and machine intelligence 11:674-693

25. Mallat SG (1989b) Multifrequency channel decompositions of images and wavelet models. I.E.E.E. Transaction on acoustic, speech and signal processing 37:2091-2110

26. Mansouri B, Hess RF, Allen HA, Dakin SC (2005) Integration, segregation, and binocular combination. J Opt Soc Amer 22:38-48

27. Marcelja S (1980) Mathematical description of the responses of simple cortical cells. J Opt Soc Amer 70:1297-1300

28. Marr D (1982a) Vision. Freeman W.H. and Company, New-York, USA, pp 54

29. Marr D (1982b) Vision. Freeman W.H. and Company, New-York, USA. pp 262

30. Nielsen M, Kamavuako EN, Andersen MM, Lucas MF, Farina D (2006) Optimal wavelet for biomedical signal compression. Med Biol Eng Comput 44:561-568

31. Pang JJ, Gao F, Wu SM (2002) Relative contributions of bipolar cell and amacrine cell inputs to light responses of ON, OFF, and ON-OFF retinal ganglion cells. Vision Research 42:19-27

32. Quiroga Q, Rosso OA, Basar E, Schürmann M (2001) Wavelet entropy in event-related potentials: a new method shows ordering of EEG oscillations. Biol Cybern 84:291-299

33. Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Research 5:583-601

34. Rosso 0A, Figliola A, Creso J, Serrano E (2004) Analysis of wavelet-filtered tonic-clonic electroencephalogram recordings. Med Biol Eng Comput 42:516-523

35. Stork DG, Wilson HR (1990) Do Gabor functions provide appropriate descriptions of visual cortical receptive field? J Opt Soc Amer 7:1362-1373

36. Tikkanen PE (1999) Nonlinear wavelet and wavelet packet denoising of electrocardiogram signel. Biol Cybern 80:259-267

37. Ts'o DY, Gilbert CD (1988) The organisation of chromatic and spatial interactions in the primate striate cortex. Journal of Neuroscience 8:1712-1727

38. Wu SM, Gao F, Maple BR (2000) Functional architecture of synapses in the inner retina: segregation of visual signals by stratification of bipolar cell axon terminals. Journal of Neuroscience 20:4462-4470

39. Yang J (1992) Do Gabor functions provide appropriate descriptions of visual cortical receptive field? comment. J Opt Soc Amer 9:334-336

40. Zalevski AM, Henning GB, Hill NJ (2007) Cue combination and the effect of horizontal disparity and perspective on stereoacuity. Spatial Vision 20:107-138

41. Zetzsche C, Röhrbein F (2001) Nonlinear and extra-classical receptive field properties and the statistics of natural scenes. Network 12:331-350